The Role of Age and Gender in Educational Expansion

The South Asian Experience in the Global Context

Petra Sauer[*]

Abstract

This paper provides an integrated analysis of within- and between-group specific trends in educational inequality. Using the IIASA/VID dataset of populations by age, sex and level of education, I calculate education Gini coefficients and decompose the overall degree of educational inequality into age, sex and within-group components for 171 countries from 1970 to 2010. I analyze the relative relevance of these components for inequality reduction and investigate the distributional outcomes of educational expansion. I find that, on average, equalization between males and females, younger and older cohorts as well as within these subgroups of the population has significantly contributed to declining educational inequality over the observed sample period around the globe. But the relative role of these components fluctuates in the process of educational expansion. First, as improvements are initiated by enhancing the educational opportunities of the youth, the gap between cohorts widens in transition phases but vanishes thereafter. Second, gaps between sexes have been reduced but are predicted to widen again if either males or females are the first to enter higher education levels. To a lesser extent, this is also true for gaps within population subgroups which can be due to the ethnic background or the social and economic status of people. Instead of a Kuznets-curve relation, I thus find evidence for education inequality to evolve in waves as education expands.

Keywords: Educational attainment, Educational Inequality, Age, Gender
JEL codes: I24,I25

---

[*]Institute for Macroeconomics and Research Institute Economics of Inequality (INEQ), Vienna University of Economics and Business (WU). E-mail: `psauer@wu.ac.at`

1 Introduction

Over the past decades, education has been rapidly expanding around the globe. Educational expansion typically entails rising population shares with primary and secondary education in low- and middle- income countries. In India, for example, 60% of the population aged 20 to 39 did not have any formal education in 1970 and secondary attainment shares were negligible. By 2010, the share of unschooled people had fallen to 26% while primary and secondary attainment shares increased to 15% and 26% respectively.[1] In contrast, achieving post-secondary education levels has become the rule in high-income countries. Pronounced dynamics in the global education structure raise questions about to the distributional consequences of educational expansion. In this regard, the distribution of education is of particular interest. It not only shapes the equalizing impact of rising educational attainment on the income distribution (see among others Ballarino *et al.*, 2013; Checchi & van de Werfhorst, 2017; Cruces *et al.*, 2011), but also the extent to which human capital accumulation positively affects growth prospects of nations (Castelló & Doménech, 2002; Sauer & Zagler, 2014).

Different degrees of educational inequality across countries over time result from the extent to which policies are able to enlarge the group of people who participate in education. This is done by improving the educational opportunities of women as well as of people from disadvantaged socio-economic backgrounds. In both respects, providing for enhanced schooling prospects of the youth secures improved educational outcomes of future generations. Consequently, different educational expansion trajectories result from different magnitudes of human capital accumulation and equalization among the youth, between men and women, and between individuals of different socio-economic backgrounds within demographic groups.

Existing research shows the distribution of educational attainment within countries to become more equal as education expands (see for example López *et al.*, 1998; Fan *et al.*, 2001). Morrisson & Murtin (2013) and Castelló & Doménech (2002) demonstrate that the strong negative relation between educational inequality and average educational attainment which has been revealed in cross-country comparisons, is mechanical and due to the decline in illiteracy. The findings of Sauer & Zagler (2014) and Meschi & Scervini (2013) provide evidence that a behavioural relationship exists within countries over time. According to Meschi & Scervini (2013), educational inequality has substantially declined in the transition toward universal basic education, but expansion of post-secondary education tends to increase the degree of inequality in the distribution of educational attainment. This contrasts with the hypothesis of an educational Kuznets curve which implies inequality to rise before it declines in the process of educational expansion.

By now, studies concerned with the distribution of education have treated all individuals within countries equally. An exemption is the analysis of Crespo-Cuaresma *et al.* (2013) who investigate age-group and gender-specific distributions of education. Educational improvements typically entail rising education levels of the youth compared to the elderly and might affect males and females differently. In this paper, I contribute to the literature by providing an integrated analysis of the evolution of gaps within- and

---

[1] These figures are obtained from the Wittgenstein Centre Data Explorer.

between demographic subgroups of the population for a global panel of countries. I thus decompose overall educational inequality into age, gender and within-group components. Using a matrix algebra approach, Silber (1989) shows how to decompose the Gini index into three components, a within- and a between-group component, and a residual term. I adapt his method so that it can be applied to aggregate education data instead of individual income data. Doing so enables me to separate inequalities which are due to inequality across age groups as well as between males and females from disparities within these groups for 171 countries over the time span from 1970 to 2010 in 5-year intervals. Moreover, I provide an intuitive interpretation of the residual term which relates to the degree of within-group inequality. I descriptively analyze the evolution of the obtained indicators of educational inequality putting the experience of South Asian countries into the global context. Finally, in a panel-regression framework, I test for non-linearities in the relation between average educational attainment, the education Gini, and each of its components.

The remainder of this paper is structured as follows. Section 2 surveys existing studies on the measurement of educational inequality. Thereafter, I use matrix algebra to derive the education Gini coefficient as a measure of between-category inequality and describe its decomposition in Section 3. I present the data I use in Section 4 and discuss the results of my decomposition analysis in Section 5. In Section 6, I test for non-linearity in the relationship between average educational attainment and the obtained inequality components. Finally, Section 7 concludes.

2   Measuring Educational Inequality

In general, education has formal and informal dimensions and comprises aspects of quality. Individuals will hence differ, among other things, according to the quantity and quality of their formal education, post-school learning and experience as well as the informal knowledge existing in their social environment. It is not possible to observe and measure all aspects of peoples educational achievement, though. Even with data from individual or household surveys, one is almost always restricted to information on formal schooling careers. That is, one observes if a person did not experience any education, has attained some basic or higher schooling and one can estimate the years associated with the respective education level. From this it follows that formal schooling is a categorical rather than a continuous variable. It has a lower boundary at zero, an upper boundary given by the duration to complete tertiary education and categories which correspond to formal education levels.

Two measures have been used primarily in order to investigate the distributional dimension of education.[2] The standard deviation of schooling has been used to explore the impact of the distribution of education on income growth and poverty reduction (e.g. Birdsall and Londoño, 1997; López et al., 1998) as well as income inequality (e.g. Lam and Levison, 1991; Inter-American Development Bank, 1999). Furthermore, similar to the concept of income inequality, standard deviations were applied to test for the existence of

---

[2]Fan *et al.* (2002) and Meschi & Scervini (2013) also calculate Theil indices of educational attainment and Castelló & Doménech (2002) additionally report the distribution of education by quintiles.

an educational Kuznets curve, i.e. an inverted U-shape relation between the distribution and the average level of schooling. By relating the standard deviation of 140 countries in 2000 to average years of schooling, Fan et al. (2002), confirmed the findings of Londoño (1990) and Ram (1990) that educational inequality first increases as the average level of schooling rises, and, after reaching a peak, starts to decline.

However, the standard deviation is only a measure of absolute dispersion and is sensitive to changes in the mean. As a measure of relative inequality, the education Gini coefficient is therefore seen as a more consistent and robust measure of the distribution of education. Some earlier studies (e.g. Maas and Criel 1982, Rosthal 1978 and Sheret 1988) used schooling enrollment figures or education finance data for calculating education Gini coefficients for small samples of mostly developing countries. These data bases do not accurately reflect the existent stock of human capital, though. Enrollment ratios are flow variables that add to future stock of human capital. Even if they constitute an indicator of access to education or equality of opportunity, they do not capture the degree of inequality in educational outcomes. Due to the availability of datasets which, by reporting attainment figures for various education levels, provide a more appropriate picture of the actual distribution of education, more recent studies calculate the education Gini based on educational attainment of the concerned population. Educational attainment is typically measured by the years of schooling achieved, the percentage of individuals that have completed primary, secondary or tertiary education levels, or by peoples actual competencies (Meschi & Scervini, 2014).

Like in its application to income inequality, the education Gini coefficient is a measure of mean standardized deviations between all possible pairs of persons and lies in a range between zero and one. A value of zero indicates a perfectly equally distributed education structure, with the opposite being true for a value of one. The former case corresponds to a situation in which the whole population attains the same education level, irrespective of which. In the latter case, one person completes for example tertiary education, while the rest of the population does not attain any formal schooling. López *et al.* (1998) were the first to derive Gini coefficients for 12 countries from attainment data. Fan *et al.* (2001) provide a detailed description of the underlying methodology, calculate Education Gini's for 85 industrialized and developing countries for the period from 1960 to 1990 and relate them to average educational attainment, educational gender-gaps and real GDP per capita. They further extend the sample to 140 countries from 1960 to 2000 in their subsequent work (see Fan et al. 2002). Thereafter, their approach has been utilized for deriving a consistent indicator of the distribution of education, that can be related to income distribution (e.g. Checchi, 2000) and income growth (e.g. Castelló & Doménech, 2002 and Sauer & Zagler, 2014). Non-conforming with earlier results, plotting Gini Coefficients against average educational attainment does not support an education Kuznets curve but reveals a strong negative relation between the degree of inequality and the average level of educational attainment. However, according to Sauer & Zagler (2014) and Meschi & Scervini (2013), the relation can be non-linear if changes of educational inequality within countries over time are taken into account.

Crespo-Cuaresma *et al.* (2013) have integrated the demographic dimension into the analysis of educational inequality. They have constructed a dataset of education Gini measures by age group and gender for 175 countries from 1960 to 2010 in 5-year inter-

vals based on the first version of the IIASA/VID[3] global dataset of populations by age and sex as well as four levels of education.[4] Investigating differential trends of educational inequality within population subgroups, they show that education is more equally distributed among the youth than among the elderly, and among men than among women.

In this article, I build on and add to the analysis of Crespo-Cuaresma *et al.* (2013) and Meschi & Scervini (2013). First, I use the most recent version of the IIASA/VID education dataset which provides a breakdown of populations into six education categories, incorporating uncompleted primary as well as lower secondary schooling (see Section 4). Second, I am able to test for non-linearities in the level-inequality relationship using a global sample of 171 countries with a longitudinal dimension. Most importantly, I decompose the overall education Gini coefficient in order to examine the relative relevance of inequality within and between age groups and gender simultaneously.

3 Matrix Algebra and the Education Gini

In an early work Silber (1989) presents a matrix approach to the computation of the Gini index of income inequality. In the following, I demonstrate how this method is adapted to the categorial structure of aggregate education data. In particular, I derive the Gini coefficient of educational attainment in matrix notation as a measure of between-category inequality. Using matrix algebra also enables to decompose the Gini index by population subgroups into three components: a within-group component equal to the weighted sum of within-group inequality, a between-group component equal to the weighted sum of between-groups inequality, and a residual term which can be interpreted in relation to the ranking of individuals within subgroups.

3.1 The Education Gini as a Measure of Between-category Inequality

According to Silber (1989), for individual data, the Gini index of inequality can be written in matrix notation as

$$I_G^E = e'Gs \tag{1}$$

where e' is a row vector with n elements equal to $1/n$. n is the number of observed individuals. If educational attainment is measured by years of schooling, one element, s_i, of the column vector s is the share of individual i's attainment in total years of schooling ($Y_T = \sum_{i=1}^{n} Y_i$) in the concerned population. The n elements of s are sorted in descending order according to individual ranks in the education distribution, so that

$$s_1 \geq s_2 \geq ... \geq s_i \geq ... \geq s_n \tag{2}$$

The linear operator, introduced by Silber (1989), is the G-matrix; which is an $n \times n$ matrix with upper-diagonal elements g_{ij} when $j > i$ equal to -1, lower-diagonal elements when $i > j$ equal to 1 and diagonal elements when $i = j$ equal to 0.

If the available information is limited to the formal duration it takes to complete an education level, individual variation within these categories vanishes. In order to demonstrate that the education Gini is thus reduced to a measure of between-category inequality, I partition the relevant vectors as well as the G-matrix. The quantity of subvectors of e and s is given by the number of categories, c. The amount of subvector elements, in turn, depends on the number of individuals, n_h, for which h is the highest education level attained. The partitioned matrix G consists of c^2 submatrices and thus has the following form

$$
\begin{pmatrix}
G(n_1, n_1) & \dots & G(n_1, n_q) & \dots & G(n_1, n_c) \\
\vdots & \ddots & & & \vdots \\
G(n_p, n_1) & & \ddots & & \vdots \\
\vdots & & & \ddots & \vdots \\
G(n_c, n_1) & \dots & \dots & \dots & G(n_c, n_c)
\end{pmatrix}
\tag{3}
$$

The main-diagonal submatrices of dimension $n_h \times n_h$ capture within-category inequality, with zeros in their main diagonal, -1 in their upper right and 1 in the lower left triangle. Submatrices, $G(n_p, n_q)$, for which $q > p$, consist of identical elements equal to -1. If $p > q$, the elements are equal to 1. Summing over partitioned elements, the education Gini can be written as consisting of a within- and between-category component,

$$
I_G^E = \sum_{h=1}^{c} e'(n_h) G(n_h, n_h) s(n_h) + \sum_{p=1}^{c} \Big[\sum_{q \neq p}^{c} e'(n_p) G(n_p, n_q) s(n_q) \Big] = I_W^E + I_B^E
\tag{4}
$$

If no information about within-category variation is available, the within component is redundant, the overall Gini index reduces to its between-category component, i.e. $I_G^E = I_B^E$, and the degree of inequality is generally underestimated. Further inspection of I_B^E enables to show how the between-category measure of educational inequality can be computed using population shares and category averages of years of schooling.

Defining the share of category h in total years of schooling, $s_{.h} = \sum_{i=1}^{n_h} Y_{ih}/Y_T$, as well as the mean individual share of years in category h, $\bar{s}_h = s_{.h}/n_h$, the between-category contribution of one pq-element for which $q > p$ can be written as

$$
e'(n_p) G(n_p, n_q) s(n_q) = -\Big(\frac{n_p}{n}\Big) n_q \bar{s}_q
\tag{5}
$$

In turn, the contribution of an element with $p > q$ looks like

$$
e'(n_q) G(n_q, n_p) s(n_p) = \Big(\frac{n_q}{n}\Big) n_p \bar{s}_p
\tag{6}
$$

Overall between-category inequality can be written as the weighted average of its individual pq-contributions, I_{pq}^E, with weights equal to the two concerning categories'

population and years-of-schooling shares respectively:[5]

$$I_B^E = \sum_{p=1}^{c} \sum_{q>p}^{c} \frac{n_p + n_q}{n}(n_p \bar{s}_p + n_q \bar{s}_q)I_{pq}^E$$

$$= \sum_{p=1}^{c} \sum_{q>p}^{c} \left(\frac{n_q}{n}\right)n_p \bar{s}_p - \left(\frac{n_p}{n}\right)n_q \bar{s}_q \tag{7}$$

Both variants of Equation (7) use the mean of schooling years in each category in conjunction with population shares. They can thus be easily calculated based on aggregate data of educational attainment. Rearranging the second part of Equation (7) enables to obtain the familiar version of the education Gini index as a weighted sum of differences in category-specific years of schooling (see Section 2):

$$I_G^E = I_B^E = \frac{1}{\bar{Y}} \sum_{p=1}^{c} \sum_{q>p}^{c} (Y_q - Y_p)p_p p_q \tag{8}$$

where $\bar{Y} = Y_T/n$ is the mean of years of schooling in the concerned population, $Y_h = \left(\sum_{i=1}^{n_h} Y_i/n_h\right)/Y_T$ is the average duration it takes to complete education category h and $p_h = n_h/n$ is the corresponding population share.

3.2 Population Subgroups

In contrast to education categories, individuals from different sub-groups of the population cannot be ordered definitely, resulting in overlapping partitions of the education distribution. Nevertheless, the matrix approach provides an intuitive and straight forward method to decompose the education Gini index into population-subgroup contributions.

Following Silber (1989), I define an additional partitioned vector, v, which is ordered first, by subgroup averages of education attainment shares and, second, by individual attainment shares within subgroups. Thus,

$$\bar{v}_1 \geq ... \geq \bar{v}_j \geq ... \geq \bar{v}_g \tag{9}$$

and

$$v_{1,j} \geq v_{2,j} \geq ... \geq v_{n_j,j} \quad \forall j \tag{10}$$

where $v_{i,j}$ is the educational attainment share of individual i in group j and $\bar{v}_j$ is the groups' mean attainment share. The number of subvectors depends on the number of groups g, and the quantity of elements, n_j, varies according to the group's population size. If also e and G are partitioned by population subgroup and defined just as in Section 3.1, a modified inequality index, $e'Gv$, can be decomposed into a within-group and a between-group component,

$$\sum_{j=1}^{g} e'(n_j)G(n_j,n_j)v(n_j) + \sum_{a=1}^{g} \left[\sum_{b \neq a}^{g} e'(n_a)G(n_a,n_b)v(n_b)\right] = I_W^E + I_B^E \tag{11}$$

---

[5]See Appendix A.

Rewriting the within-group elements in terms of group attainment and population shares enables to obtain the within-group component as a weighted average of within-group inequality, with weights equal to the groups' population and attainment shares $(v_{.j})$ respectively. Thus,

$$I_W^E = \sum_{j=1}^{g} \frac{n_j}{n} v_{.j} I_{Gj}^E \tag{12}$$

where within-group inequality, I_{Gj}^E, is the between-category Gini index of subgroup j. Beyond that, in a similar manner as in Section 3.1, the between-group component of the education Gini can be shown to be equal to a weighted average of pairwise contributions, I_{ab}^E, with weights equal to the two groups' joint population and attainment shares:

$$I_B^E = \sum_{a=1}^{g} \sum_{b>a}^{g} \frac{n_a + n_b}{n} (n_a \bar{v}_a + n_b \bar{v}_b) I_{ab}^E$$
$$= \sum_{a=1}^{g} \sum_{b>a}^{g} \left(\frac{n_b}{n}\right) n_a \bar{v}_a - \left(\frac{n_a}{n}\right) n_b \bar{v}_b \tag{13}$$

Defining mean years of schooling in subgroup j as $\bar{Y}_j = \sum_{i=1}^{n_j} Y_{ij} / n_j$, the second part of Equation (13) can be rearranged in order to obtain the between-group contribution as a weighted sum of differences in subgroups' mean years of schooling:

$$I_B^E = \frac{1}{\bar{Y}} \sum_{a=1}^{g} \sum_{b>a}^{g} (\bar{Y}_a - \bar{Y}_b) p_a p_b \tag{14}$$

Finally, the difference between the inequality measures obtained from using the definitely ordered versus the reordered attainment share vectors, $e'Gd = e'(s - v) = e'Gs - e'Gv$, builds the third component of the Gini index decomposition. This factor can be interpreted as the intensity of modifications necessary to rank individuals according to their groups' educational attainment, or, as the degree to which groups are overlapping.

4 Data

In order to compute education Gini coefficients according to Equation (8) and decompose them by age and gender, I require information about the educational structure of populations and the typical years people spend in school to attain an educational level. I obtain the full distribution of educational attainment over six categories by five-year age groups and gender from data provided by the Wittgenstein Centre for Demography and Global Human Capital (WIC)[6]. To assemble this dataset, Goujon *et al.* (2016) collect baseline data for 2010 from censuses if available, or from surveys otherwise, and apply harmonization procedures to obtain six education categories which correspond to UNESCO's International Standard Classification of Education (ISCED 1997) as summarized in Table

---

[6]This most recent version of the IIASA/VID education dataset is publicly available via `http://dataexplorer.wittgensteincentre.org/shiny/wic/`.

Table 1: Categories of educational attainment

	Category[a]	ISCED 1997 level
1	No education	No level of ISCED0
		Grade 1 of ISCED 1 not completed
2	Incomplete Primary	Incomplete ISCED1
3	Primary	Completed ISCED1
		Incomplete ISCED2
4	Lower Secondary	Completed ISCED2
		Incomplete ISCED3
5	Upper Secondary	Completed ISCED3
		Incomplete ISCED4 or 5B
6	Post-secondary	ISCED 4 & 5B[b], ISCED 5A & 6 [c]

[a]See Goujon *et al.* (2016).

[b]First diploma, shorter post-secondary courses

[c]Longer post-secondary courses, post-graduate level

1. Goujon *et al.* (2016) subsequently adopt the demographic method of multistate projection using education-specific estimates of fertility, mortality and migration in order to project from 2010 backward to 1970 and forward to 2060. I use the historical part of their data to compute educational attainment shares for the total population aged 15, each of 18 5-year age groups, males and females.

As shown in Table 1, education categories 3 to 5 are composed of individuals who complete the respective but not the subsequent education level. For example, people in category 3 not only go through primary but also through some lower secondary education. Potancoková *et al.* (2014) adjust data on formal schooling cycles from the UNESCO Institute for Statistics (UIS) to allow for incomplete levels using region-specific correction factors based on survey estimates. They also provide estimates of the average years of incomplete primary education but apply a constant of four years for post-secondary education to balance the diverseness of this category. I use their adjusted country- and cohort-specific duration data to measure the typical duration of each education level in the computation of average years of schooling and education Gini coefficients (see Equation 8).

I obtain a dataset with information on average educational attainment and educational inequality for the total population aged 15 and over as well as for each of 18 five-year age groups, males and females. I cover 171 countries over the time period from 1970 to 2010 in five-year intervals. Figure 1 presents the resulting structure for each country-year observation of the data for the example of India in 2010. The average educational attainment of 5.49 years of schooling and the corresponding education Gini, equal to 0.54, mask differences between demographic subgroups. Plotting both measures against age gives rise to downward and upward sloping curves for the average level and inequality respectively. This indicates that educational improvements start among the youth. If

these not only affect the average level but also the spread, education is more equally distributed among the youth than among the elderly. Even if these relations holds globally, countries differ with respect to the slopes, i.e. the magnitude of educational expansion and equalization.[7] Moreover, they differ with regards to the gender-gap which exists not only for the average level of educational attainment but also for its distribution. In most countries, education is more equally distributed among men than among women, at least for older cohorts. The polarization between a small segment of highly educated and a broad group with basic education is thus more pronounced among women than among men. Differences between men and women tend to be lower or vanishing for younger cohorts. However, inequalities below the age of 25 can be substantially underestimated as some individuals have not completed post-secondary education. In India, the gender-gap in educational inequality is relatively large and constant at two Gini points for cohorts older than 45.
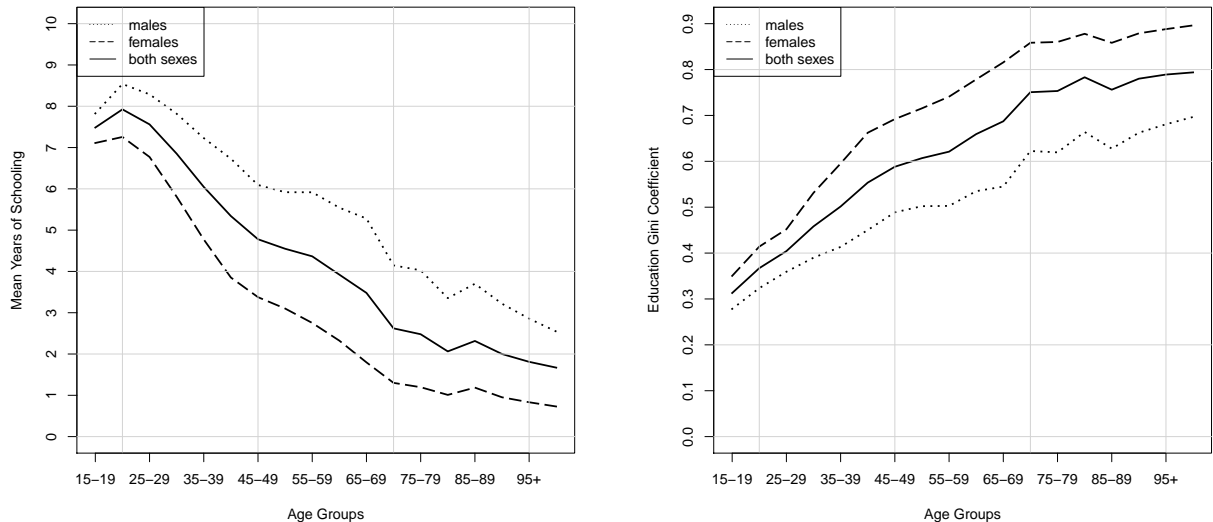

Figure 1: Average Attainment and Educational Inequality, India 2010

In order to ascertain to measure the degree of inequality in completed educational attainment, I restrict my analysis to cohorts aged 25 and over. The finer the grouping, the more homogeneous groups are. This increases the relative importance of between-versus within-group variation. Considering this impact on the Gini index decomposition, I construct three broader age groups for people aged 25 to 44, 45 to 64 as well as 65 and over. I thus obtain 6 subgroups of the population which I use to examine the relevance of age, gender and within-group differences for the degree of inequality in educational attainment.

[7]See Crespo-Cuaresma *et al.* (2013) for a detailed discussion and analysis of the dynamics of age-group specific education inequality.

5 Decomposing Educational Inequality - South Asia in the Global Context

Concerning the distribution of educational attainment, South Asian (SA) countries are among the most unequal of the world. The average education Gini of the region is not only higher than the global average in 2010 (0.33), but also larger than the respective value for Sub-Saharan Africa (SSA) (see Table 2). Both regions also report the lowest number of mean years of schooling. In contrast, high income (HI) OECD as well as Central Asian and Eastern European (CAE) countries show the lowest Gini values in conjunction with the highest average level of educational attainment, on average. The latter region was able to reduce educational inequality by 63% between 1970 and 2010, what was mainly driven by the significant reduction in educational inequality in the South Eastern part of Europe.[8] Substantial improvements also took place in East Asia and the Pacific (EAP), which is in the medium spectrum of global educational inequality. With a change of 52%, the general trend towards more equal education distributions was particularly pronounced in Middle Eastern and North African (MENA) countries. The decline over time is not only sluggish in HI OECD countries, where the initial level was already low, but also in persistently high-inequality countries in SA and SSA. Also Latin American and Carribean (LAC) countries remain at a relatively high level of educational inequality.

Table 2: Summary Statistics 1

	$\bar{Y}^a$	$\Delta\bar{Y}^b$	I_G^E	ΔI_G^E
CAE	10.52	77.60	0.16	-62.50
HI OECD	11.63	36.82	0.16	-34.06
EAP	8.24	127.85	0.29	-47.64
LAC	8.36	84.10	0.30	-38.06
MENA	7.89	247.52	0.37	-52.38
SSA	4.79	257.93	0.54	-33.76
SA	4.36	259.16	0.59	-30.17
Global	8.21	91.56	0.33	-41.04

[a]Levels are from 2010.
[b]Total Change between 1970 and 2010 in %.

---

[8]The group of CAE countries consists of countries in Eastern Europe (e.g. Latvia and Ukraine), in Central Asia (e.g. Armenia and Kazakhstan) as well as of countries in South-Eastern Europe (e.g. Albania, Macedonia and Turkey). While educational inequality is historically relatively low in the former two groups of countries, the latter group had relatively high levels of educational inequality in the 1970s and 1980s and was able to reduce it significantly thereafter.

In accordance with the existing literature, cross-country comparisons indicate a negative relation between average educational attainment and educational inequality. However, Sauer & Zagler (2014) and Meschi & Scervini (2013) have argued that this needs not be true for changes within countries over time. Moreover, Crespo-Cuaresma *et al.* (2013) have shown that trends in the overall degree of educational inequality can hide differences in inequality within age groups and genders. Beyond that, trends in between-group inequality matter for the distributional outcomes of educational expansion. Differences in inequality reduction might stem from the extent to which countries secure equalization between males and females as well as between socio-economic groups. But educational expansion can enlarge gaps between younger and older cohorts. In order to provide an integrated picture of differences between age groups and sexes, as well as differences within population subgroups, I apply equations (8), (12) and (14) to decompose the education Gini coefficient of the total population aged 25 and over of 171 countries from 1970 to 2010 into five components as follows,

$$I_G^E = I_{age}^E + I_{age/sex}^E + I_{sex}^E + I_{within}^E + I_{residual}^E \tag{15}$$

where I_{age}^E captures the contribution of gender-specific differentials within age-groups while $I_{age/sex}^E$ compares the educational attainment of different sexes *and* cohorts. I_{sex}^E is the component of the education Gini which is due to differentials between males and females of the identical age groups. The extent of inequality within population subgroups, in turn, is given by I_{within}^E. Finally, $I_{residual}^E$ is the residual component.

The upper panel of Figure 2 shows how educational inequality and each component evolved over time in South Asian countries. It reveals that the decline in overall inequality involves declining between- and within group inequality.[9] However, the relative relevance of age, gender and within-group inequality varies over time and across countries. The lower panel of Figure 2 plots each component's share in total inequality over time. In general, the reduction of differences between females and males has been the most significant factor which contributes to declining educational inequality. Nevertheless, the contribution of gender inequality is not only relatively large (greater than 20%) until the 1990s in Bhutan, Nepal and Pakistan, but also did not decrease substantially before the end of the 20th century. The slow decline in overall educational inequality in these countries is thus partly explained by the sluggish reduction in gender inequality. Since the beginning of the new millennium, gender inequality has improved significantly in Bhutan and Nepal. In India, educational gaps between males and females made up a constantly declining share in total inequality (from 16% in 1970 to 9% in 2010). Such a development was particularly pronounced in Bangladesh where the respective ratio declined from 19% in 1970 to 5% in 2010. In contrast, the relative relevance of gender inequality has already been low since the 1970s in the Maldives but decreased further by 4 percentage points to 1.5% in 2010.

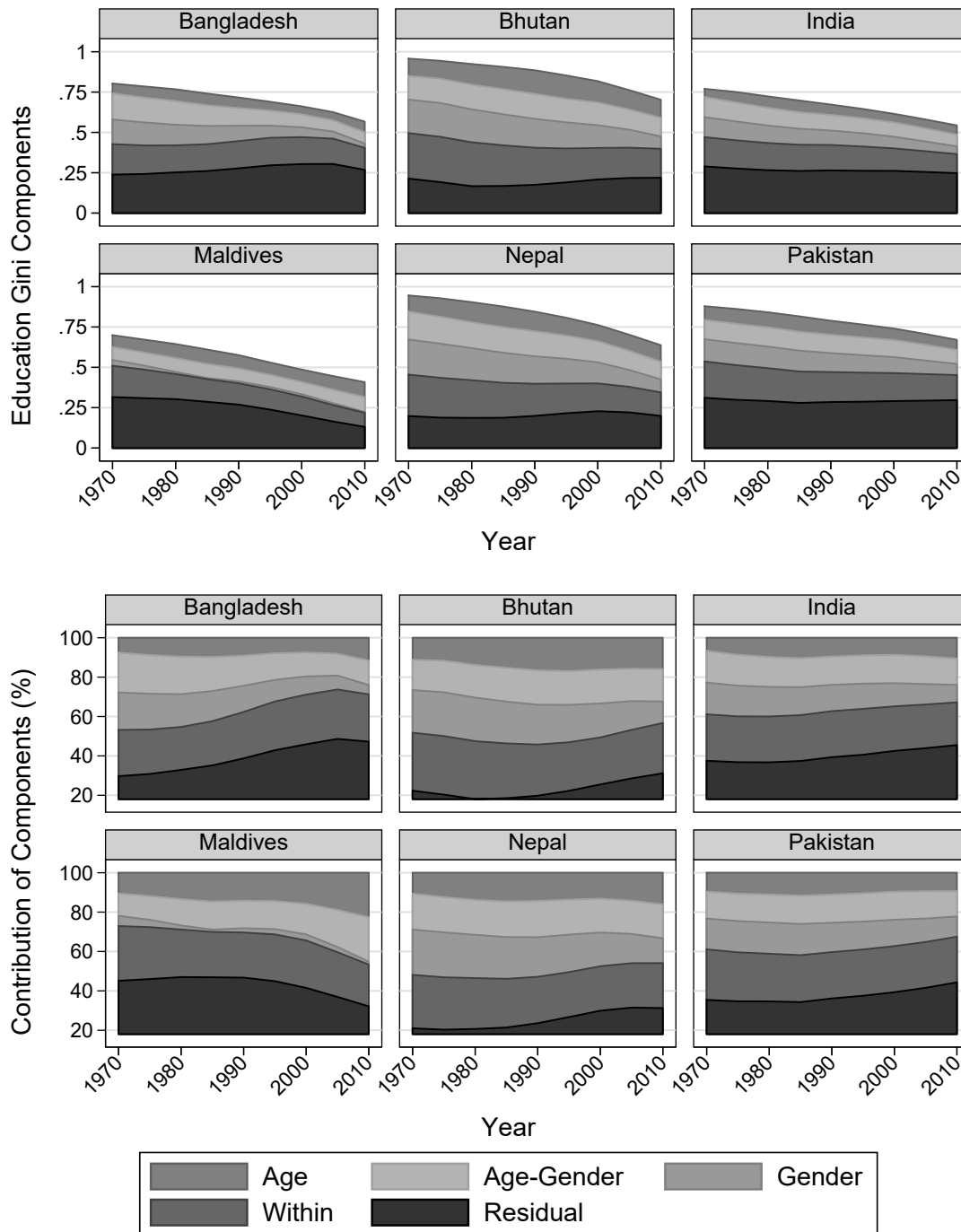[9]See below for a discussion about the residual term.

Figure 2: Educational Inequality Components in South Asia

The contribution of within-group inequality is relatively constant around the regional average of 24% in each country. Even if I observe a significant declining time trend of the within-group contribution for the whole region, in neither country did the share fall below 20% over the observed time period. This implies that educational divides within demographic subgroups of the population continuously exert a significant contribution to the total level of educational inequality. As educational improvements start among the youth, gaps between age groups widen in transition phases. Thus, the relative contribution of both age components increases as the relative relevance of gender and within-group inequality declines, particularly in the Maldives, but also more recently in Bangladesh, Bhutan, India and Nepal.

The developments in South Asian countries already hint to more general findings concerning the the relationship of within- and between-group inequality with overall educational inequality and educational expansion. To put the experience of South Asia into the global context, I compare world regions in Table 3 and investigate the relation between the education Gini of the total population aged 25 and over with each of its components using the global sample of countries in Figure 3.

Table 3: Summary Statistics: Components

		I^E_{age}	$I^E_{age/sex}$	I^E_{ageF}	I^E_{ageM}	I^E_{sex}	I^E_{within}	$I^E_{residual}$
HIOECD	Level	0.03	0.03	0.02	0.01	0.00	0.03	0.08
	%	15.86	16.04	10.85	5.02	2.58	16.11	49.40
CAE		0.03	0.03	0.02	0.01	0.01	0.02	0.07
		17.52	17.39	13.02	4.50	3.65	16.13	45.31
EAP		0.05	0.05	0.03	0.02	0.01	0.06	0.13
		16.73	17.30	10.48	6.26	3.58	19.22	43.16
LAC		0.04	0.04	0.02	0.02	0.01	0.06	0.15
		13.75	13.41	8.62	5.13	2.61	19.38	50.86
MENA		0.05	0.05	0.03	0.02	0.02	0.09	0.16
		13.32	13.81	7.98	5.34	5.95	25.16	41.75
SSA		0.06	0.08	0.04	0.03	0.05	0.13	0.22
		13.00	15.08	7.63	5.37	8.66	23.08	40.19
SA		0.08	0.09	0.04	0.04	0.05	0.14	0.23
		14.34	15.81	7.33	7.01	8.16	23.09	38.61

The age component tends to be higher the larger the increase in educational attainment of the youth as compared to the elderly. Differences across age groups of same sexes are thus closely connected to educational improvements. The contribution of the age component is thus lowest in the high-inequality countries of SSA (13%) in 2010, where

younger cohorts tend to be as similarly low educated as their predecessors. The relative importance of age is largest in CAE countries (18%), again due to the high dynamic countries in South-Eastern Europe, as well as in EAP (17%), where e.g. South Korea experienced a period of substantial educational expansion across all education levels. In HI OECD countries differences across age groups result from the expansion of post-secondary education. While SA and MENA have consistently enlarged access to basic education over the last decades, LAC countries show a persistent education structure at medium levels, resulting in a relatively low contribution of the age component.

Generally, the upper-left panel of Figure 3 depicts a negative relation between the relative relevance of the age component and the total education Gini. However, closer inspection of the underlying processes hint to repeating inverted U-shaped relations between the relevance of age-differences and educational inequality. First, while the relative contribution of age increased over the whole time span in SA, it increased until 1990, but has been decreasing since then in all other world regions.[10] This implies that the attainment shares of younger cohorts increased relative to their predecessors before they started to stagnate. Second, not only are age-group differences substantially more relevant for females than for males, but they also exhibit opposing relations with overall educational inequality (see upper-right panel of Figure 3). The positive relation for males is due to converging education levels of successive male generations, indicating their relative advance in the process of educational improvements. In contrast, the attainment share of young female cohorts is continuously higher than that of their predecessors over the whole time span, resulting in a downward sloping line. The gender and time-specific patterns of the age component highlight the varying role of the distribution of education between cohorts in the process of education expansion, with an increasing divide if younger generations become higher educated than the older and a narrowing gap as these cohorts age. Beyond that, these findings indicate that before including females, education expansion started among young males.

The contribution of differences between age groups and gender is almost equal to the gender-specific age component in all regions except SSA and SA. The divide in educational attainment is thus larger, comparing, for example, males and females aged 65 and 25 respectively, than comparing females of the concerning age groups. As Figure 4 shows for South Asian countries, the share of women aged 45 to 64 in total educational attainment is not only lower than that of younger women, but also significantly lower than that of younger males. This difference is in turn larger than that which compares the attainment shares of medium aged men and young women. The insignificant relation between the age/sex component and overall inequality again hides differences between males and females. While increasing gaps between young women and older men contribute to lower educational inequality, the reverse is true for gaps between young males and older female cohorts. There is thus an additional factor of gender to age-group inequalities in education.

Closing the divide between males and females of same cohorts has significantly con-

---

[10]In HI OECD, LAC, SA and SSA even the absolute level of the age component significantly increased until 1990. These results stem from fixed-effects regressions of each component or its relative contribution on time and are available from the author upon request.
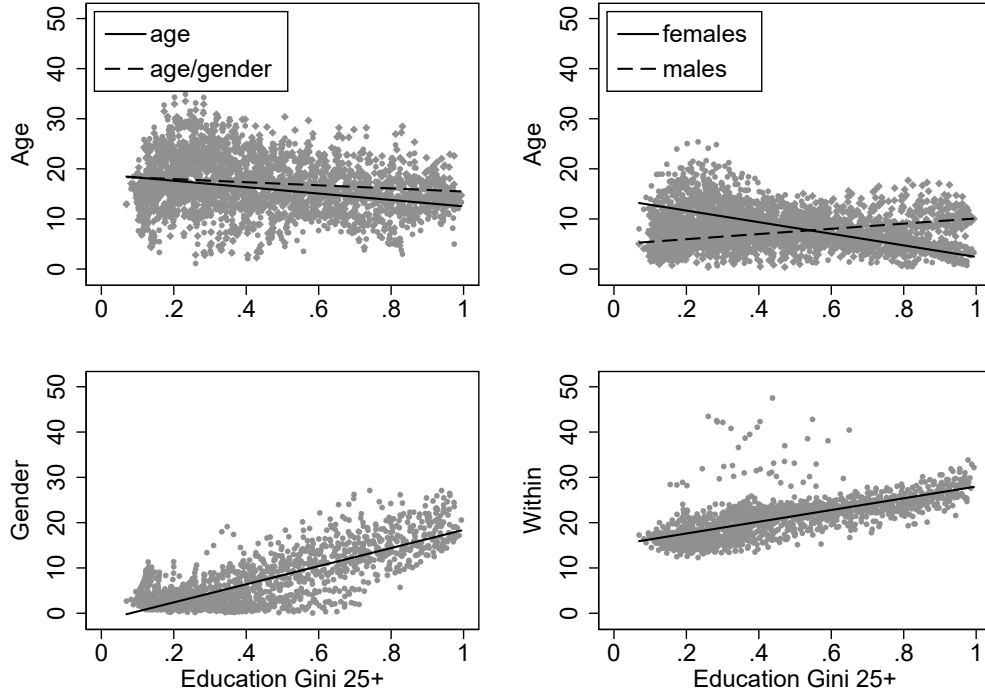
Figure 3: Contribution of Components (%) & Overall Inequality

tributed to declining educational inequality around the globe. Not only do I observe a positive relation between the total Gini index and the relative contribution of the gender component, but also a decreasing trend of its contribution over time for all world regions. In MENA, SSA, SA and EAP, this is due to continuously increasing education shares of young women who follow up in completing basic education levels. In 82 countries the education share of young females was even higher than that of their male counterparts in 2010. This is true for some countries in LAC and CAE where young-female shares have been constantly higher than those of young males (e.g. Georgia, Uruguay), but mainly for countries where the switch took place in the 1990s (e.g. Hongkong, Brazil, Albania). In HI OECD countries, the education shares of both genders evolve almost simultaneously over the whole sample period, resulting in a constantly low contribution of the gender component. Compared to its low overall education Gini, CAE shows a relatively high contribution of gender inequality as gender differences in South Eastern Europe have been relatively large. On the other hand, the experience of LAC countries shows that gender differences can be marginal even if overall educational inequality is relatively high. From this it follows that differential institutions are a relevant factor in shaping gender inequality.

Educational divides within subgroups of the population arise from segmentation along various social lines, including ethnicity, geography and divisions between individuals from different socio-economic backgrounds, among other things. The distribution of education
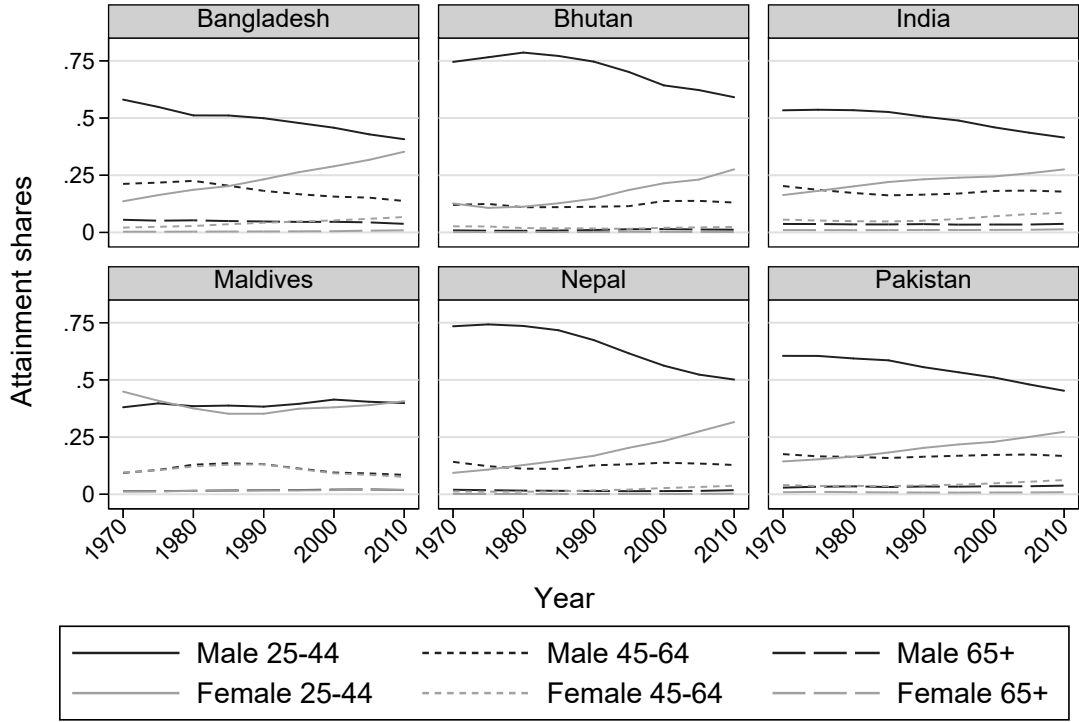
16

Figure 4: Attainment shares ($\bar{v}_j$) by Age Group and Gender

within population subgroups thus plays a significant role in shaping the overall level of inequality. Besides the degree of within-group inequality, the within-group component depends on the each group's share in the population as well as in total educational attainment. It is thus generally larger for females than for males, and for older[11] than for younger cohorts. Even if its relevance is generally increasing in the level of education inequality, inequality within population subgroups makes up the largest share in educational inequality but is only marginally decreasing over time.[12] The relative within-group component contribution is greater than all three between-group components in LAC, EAP, MENA, SSA and SA. Notably, with a relative contribution of 25% in 2010, the within-component plays a considerable role in countries of the Middle East and North Africa, compared to the level of educational inequality in the region.

The within-group component compares individuals of one group while the between-group component compares population groups by their average educational attainment. The residual component, in turn, interacts with both factors. First, the lower between-group inequality, the more are subgroups of the population overlapping. Or, the more permutations are necessary to rank individuals first, by the average educational attainment of their group and second, by the relative position within their group. Second,

---

[11]This effect can be negligible if life expectancy is very low, so is the population share of the elderly.

[12]Globally, the relative contribution of within-group inequality declined by 0.05 percentage points each year. In comparison, the respective time trend for the gender component is 0.13.

comparing individuals from different groups solely by group averages abstracts from comparisons of outliers not represented by their group's mean attainment. The more polarized groups are, i.e. the greater the spread within groups, the more relevant these comparisons become. Hence, the higher the residual component should be.

The residual component explains a large part of educational inequality in each world region (see upper panel of Figure 5). While I find its absolute level to decrease with the education Gini, its relative contribution rises as overall educational inequality declines (see Figure 5). In LAC, particularly the relevance of gaps between males and females is marginal, resulting in a relatively large residual contribution. As differences between population groups are still relevant, its contribution of 39% and 40% is relatively low in SSA and SA respectively. Compared to their medium overall degree of inequality, countries in the Middle East and North Africa stand out with their low level of residual contribution. This is due to outliers in the Middle East[13], characterized by a low between-group-inequality contribution in conjunction with particularly high within-group inequality. Even if this suggests a negative and positive relation of residual inequality with between- and within-group inequality respectively, the lower panel of Figure 5 does not provide unambiguous evidence. The residual component approaches zero as gaps within and between groups vanish. It also tends to be low for high values of between-group inequality. In between the extremes, the variation of the residual component is large. Plotting the residual against the within-group component provides indication that this can be due to countries differing with respect to within-group inequality (see Section section:expansion).[14]

To sum up, the analyzed components of educational inequality exhibit differential roles across regions and vary over time in the process of educational improvements. Even if the sample period I observe is dominated by equalizing processes between cohorts, differing trends between males and females as well as over time indicate that the gap between age groups fluctuates in conjunction with the level of educational attainment. This is also true for differences between sexes. My findings indicate that educational expansion processes at basic levels have predominantly started among young males before including females. However, women outperform men at higher education levels. Nevertheless, closing the gap between genders of equal age groups has significantly contributed to the declining trend of overall education inequality. The contribution of within-group inequality, in turn, is relatively large and only marginally decreasing over time. Beyond that, polarization within groups impairs between-group comparisons, resulting in larger residual components of education inequality.

---

[13]Especially Quatar and the United Arab Emirates.

[14]This relation is even more visible if I exclude MENA countries with particularly high within-group contributions but low shares of between-group components.
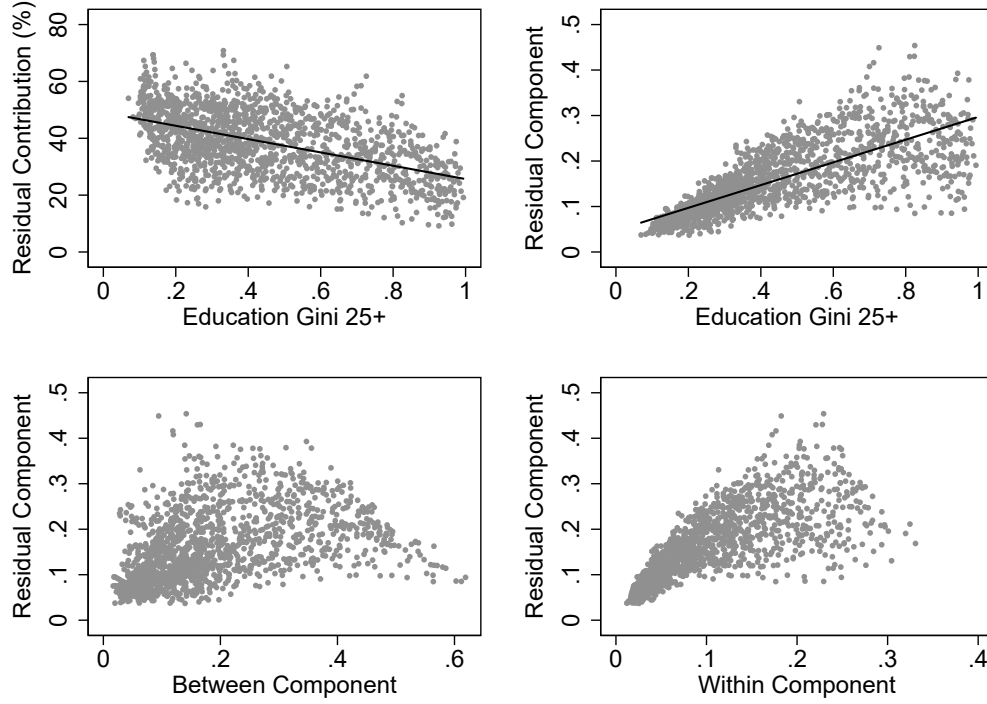
18

Figure 5: Between - Within & Residual Component

6 The Distributional Impact of Educational Expansion

Convergence of educational attainment within and between age groups as well as sexes has accounted for declining inequality in the distribution of education. However, to what extent has educational expansion around the globe contributed to this trend? Moreover, in how far have components of educational inequality have been affected differently?

Previous work on the distribution of education found evidence for a strong negative relation between the degree of educational inequality, measured by the education Gini coefficient, and average educational attainment, mostly measured by mean years of schooling. Yet, Sauer & Zagler (2012) have observed that even if this relation holds across countries, it need not be strictly negative within countries over time. The evidence of Meschi & Scervini (2013) provides support for an U-shaped relation between the average level and the distribution of education. Accordingly, educational inequality has substantially declined as education expanded, but shifting education structures towards post-secondary levels tend to increase educational inequality. Their findings contrast with the hypothesis of an educational Kuznets curve, which implies an *inverted* U-shape, i.e. adverse followed by favourable distributional consequences of educational expansion.

Also my findings in Section 5 suggest a non-linear relation between educational in-

Table 4: Regression Results: Components

	I_G^E	I_{age}^E	$I_{age/sex}^E$	I_{Fage}^E	I_{Mage}^E	I_{sex}^E	I_{within}^E	$I_{residual}^E$
$\bar{Y}$	-0.106***	-0.014***	-0.014***	0.004***	-0.018***	-0.031***	-0.030***	-0.049***
	(0.002)	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)	(0.001)	(0.002)
$\bar{Y}^2$	0.003***	0.000***	0.000***	-0.001***	0.001***	0.002***	0.001***	0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
dur^C	-0.063***	-0.018*	0.008	-0.002	-0.016**	0.022**	-0.006	-0.062***
	(0.017)	(0.009)	(0.010)	(0.005)	(0.007)	(0.011)	(0.007)	(0.011)
I_B^E								-0.606***
								(0.013)
I_W^E								0.130***
								(0.046)
$Cons$	1.491***	0.296***	0.066	0.051	0.244***	-0.052	0.279***	1.054***
	(0.157)	(0.084)	(0.091)	(0.045)	(0.063)	(0.096)	(0.060)	(0.103)
Obs	1,539	1,539	1,539	1,539	1,539	1,539	1,539	1,539
R^2	0.942	0.621	0.640	0.459	0.698	0.715	0.881	0.778
N	171	171	171	171	171	171	171	171

Standard errors in parenthesis, * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

equality and average attainment which is driven by the evolution/simultaneous responses of between- and within-group inequality. I test whether a non-linear relation between each component and average educational attainment exists in a panel regression framework. Beyond that, I investigate the distributional impact of the education policy to increase compulsory education and test for the presumed negative and positive relation of the residual component with between- and within-group inequality respectively. A simple model which aims to explain the level as well as the relative contribution of each component can be written as,

$$\text{Comp}_{i,t} = \alpha_i + \beta_1 \bar{Y}_{i,t} + \beta_2 \bar{Y}_{i,t}^2 + \beta_3 dur_{i,t}^C + \lambda_t + \varepsilon_{i,t} \tag{16}$$

where $\text{Comp}_{i,t}$ is the concerning component of educational inequality or its relative contribution. $\bar{Y}_{i,t}$ is a measure of mean years of schooling in the population aged 25 and over. Its square accounts for the presumed non-linearity in the relation of interest. $dur_{i,t}^C$ is the formal duration necessary to complete the lower secondary level, which I use as a proxy for compulsory education. This data is taken from UIS (see Section 4) and measured as an average over the years each of 16 cohorts above 25 went to school to complete lower secondary education. Thereby, the time lag between the education policy and educational outcomes is accounted for. Only in the regressions with the residual component as dependent variable, I also include the between- (I_B^E)- and the within-group (I_W^E) component as regressors. I include country- (α_i) and time-specific (λ_t) intercepts. Finally, $\varepsilon_{i,t}$ captures the time varying component of the error term. I estimate the parameters in Equation 16 using a fixed-effects estimator as the expected value of the time-invariant country-specific effects conditional on the explanatory variables cannot be assumed to be zero. The results are provided in Tables 4 and 5.

Table 5: Regression Results: Contribution of Components

	I_{age}^E	$I_{age/sex}^E$	I_{Fage}^E	I_{Mage}^E	I_{sex}^E	I_{within}^E	$I_{residual}^E$
$\bar{Y}$	1.339***	2.045***	1.970***	-0.632***	-2.565***	-0.767***	0.000
	(0.163)	(0.161)	(0.097)	(0.096)	(0.126)	(0.084)	(0.000)
$\bar{Y}^2$	-0.064***	-0.071***	-0.090***	0.026***	0.142***	0.009*	-0.000
	(0.009)	(0.009)	(0.006)	(0.006)	(0.007)	(0.005)	(0.000)
dur^C	-0.229	3.105*	0.711	-0.939	4.957***	3.410***	0.000
	(1.770)	(1.757)	(1.059)	(1.050)	(1.366)	(0.912)	(0.000)
I_B^E							-1.000***
							(0.000)
I_W^E							-1.000***
							(0.000)
$Cons$	12.651	-18.143	-5.633	18.284*	-27.992**	-6.049	100.000***
	(16.081)	(15.964)	(9.625)	(9.539)	(12.417)	(8.283)	(0.000)
Obs	1,539	1,539	1,539	1,539	1,539	1,539	1,539
R^2	0.222	0.255	0.478	0.453	0.580	0.285	1.000
N	171	171	171	171	171	171	171

Standard errors in parenthesis, * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table 4 presents the results using the absolute levels of the education Gini and its components as dependent variables. The total education Gini of the population aged 25 and over is predicted to significantly decrease in the process of educational expansion until 17 years of schooling, but to increase thereafter. The non-linear relation is less pronounced for both age components. The part which captures differences between same-sex cohorts turns insignificant at a level (33 years) well above the sample range of mean years of schooling. Columns (3) and (4) reveal that the small impact is due to different effects for males and females. While differences between female cohorts increase as education expands up to three years, male cohorts converge until 11 years of schooling. The finding that the education distribution among males is increasingly persistent across generations provides additional evidence that young males have been the first to benefit from improvements in the transition to universal basic education. In contrast, young women tend to outperform young men at higher education levels. Gaps between males and females are thus predicted to increase as of 8 years of schooling. At the average level of education in SA (see Table 2), an increase by one year *reduces* the gender component by 31% compared to its 2010 level (see Table 3). Conversely, it *increases* the 2010 level of gender inequality in HI OECD countries by a factor of 3.75. Even if the estimated relationship between average educational attainment and within-group inequality equally follows a U-shape, the turning point at 13 years of schooling is higher, so that only more recent observations from Canada, Finland and Germany fall above that level. In contrast, SA countries would be able to reduce their level of within-group inequality by 15% in relation to its 2010 level, as average attainment increases by one year.

Looking at relative contributions of educational inequality components (Table 5) re-

veals a somewhat different picture. Most importantly, I find evidence for an inverted U-shape of the relation between the relative contribution of both age related components and mean years of schooling. Thus, as education expands, differentials between age groups become increasingly important, but their relevance for overall inequality vanishes as educational attainment of successive young cohorts stagnates. Columns (3) and (4) show that the observed educational Kuznets curve w.r.t. age is driven by the evolution of the distribution of education among females. The relative relevance of gaps between women of different age cohorts increases as education expands up to 8 years of schooling and declines thereafter, while the reverse is true for gaps between male cohorts. In turn, the relative relevance of generational differences between sexes increases up to 12 years of schooling, but is relatively constant thereafter. The estimated relation between average educational attainment and the relative contribution of gender- and within-group inequality resembles that obtained from looking at the absolute levels of components. While the turning point is equal for the gender component, it is much later, at 24 years of schooling for the within-group component. For the latter, the estimated effects in both directions are small, what indicates that the relevance of within-group inequality is relatively constant in the process of educational expansion.

Increasing the formal duration of compulsory schooling significantly contributes to declining educational inequality via reducing generational gaps between males. However, as they are the first to benefit from the education policy, differences between (young) males and females significantly increase. Moreover, extending the duration of compulsory education significantly increases the absolute level as well as the relative contribution of gender inequality. Also gaps within population subgroups and between age-groups of different sexes become more relevant. The broader provision of basic education to young males thus contributes to the generation of educational gender gaps.

The results in the last Column of Table 4 reveal the presumed relations between within-, between-, and residual inequality, described in Section 5. The residual component tends to be low if between-group inequality is large, and high if education is unequally distributed within population subgroups. Taken together, these findings can explain the high variation of the residual component with increasing between- and within-group inequality respectively (see lower panel of Figure 5). Interestingly, average years of schooling and the formal duration of compulsory schooling exert an impact on residual inequality which is additional to their effects via between- and within-group inequality. In contrast, changes in the relative relevance of the residual component are fully explained by changes in the relative contribution of the other Gini components (see Table 5).

7 Summary & Conclusions

This paper provides an integrated analysis of the evolution of educational inequality within and between demographic subgroups of the population in the process of educational expansion. For 171 countries over the time span from 1970 to 2010, I therefore decomposed the education Gini coefficient of the total population aged 25 and over into components which measure differences in average educational attainment between age groups, males and females, and a within-group component. Investigating these new indicators provides

insights into education expansion trajectories around the globe.

Educational inequality between age groups, males and females, and between individuals from different socio-economic backgrounds within demographic groups exhibit differential roles across world regions and vary over time as societies become educated. In general, my results provide indication for educational inequality to change in waves with educational attainment. As long as specific groups are the first to benefit from improvements, inequality first rises but decreases as larger parts of the population take part in education, and remains constant until further advancements again exert disequalizing effects. Over the sample period I observe, educational inequality has substantially declined, but tends to increase as education expands further. This U-shaped relation is in line with previous evidence provided by Meschi & Scervini (2013) and contradicts the hypothesis of an educational Kuznets curve.

Each component contributes differently to the overall trend. The relative relevance of divides between age groups increases as younger generations become higher educated and decrease as these cohorts age. I find evidence on gender-specific cohort effects which indicate that the expansion of basic educational levels has predominantly started among young males before including females. Yet, educational expansion has significantly contributed to closing the gap between genders of equal age groups, what has in turn added to the declining trend of overall educational inequality throughout the observed sample period. In contrast, young women tend to be higher educated than their male counterparts, particularly in highly educated countries. The educational gender gap is thus predicted to increase in the transition to higher education levels. Inequality in the distribution of education within population subgroups accounts for a large part in overall educational inequality and its relevance is only marginally decreasing over time. Thus, the ethnic background or the social and economic status of people continues to determine the educational prospects of people around the globe. My findings indicate that even if the transition to universal basic education has contributed to declining inequality, shifting the education structure towards post-secondary levels is not only able to increase gaps between age cohorts, but also between males and females, as well as between individuals within demographic groups. Hence, as long as societies are segmented along various lines, policies benefit particular people. A comprehensive understanding of the distribution of education between and within population groups can help to identify these target groups, thereby assisting to improve the distributional impact of policies aimed at educational expansion.

The new indicators of educational inequality between- and within demographic subgroups of the population enable further research on the causes, macro-economic consequences and broader societal effects of educational inequality. For example, the age component measures the extent of educational expansion and can be used to examine the consequences for economic growth. Relating the measures of educational gender inequality to democratization, female labor force participation or fertility can contribute to the understanding of the role of women in development processes. Moreover, they can be used to analyze whether closing education gender gaps have contributed to narrowing gender wage gaps at the aggregate level.

References

BALLARINO, GABRIELE, BRATTI, MASSIMILIANO, FILIPPIN, ANTONIO, FIORIO, CARLO, LEONARDI, MARCO, & SCERVINI, FRANCESCO. 2013. Increasing Educational Inequalities? *In:* SALVERDA, WIEMER, NOLAN, BRIAN, CHECCI, DANIELE, MARX, IVE, MCKNIGHT, ABIGAIL, TÓT, ISTVÁN GYORGY, & VAN DE WERFHOST, HERMAN (eds), *Changing Inequalities in Rich Countries.* Oxford: Oxford University Press.

BARRO, ROBERT J., & LEE, JONG-WHA. 2010. *A New Dataset of Educational Attainment in the World, 1950-2010.* NBER Working Paper 15902.

BENAABDELAALI, WAIL, HANCHANE, SAID, & KAMAL, ABDELHAK. 2012. Chapter 13 Educational Inequality in the World, 1950-2010: Estimates from a New Dataset. *In:* BISHOP, JOHN A., & SALAS, RAFAEL (eds), *Inequality, Mobility and Segregation: Essays in Honor of Jacques Silber (Research on Economic Inequality, Volume 20.* Emerald Group Publishing Limited.

BIRDSALL, NANCY, & LONDONO, JUAN LUIS. 1997. Asset Inequality Matters: An Assessment of the World Bank's Approach to Poverty Reduction. *American Economic Review,* **87**(2), 32–37.

CASTELLÓ, AMPARO, & DOMÉNECH, RAFAEL. 2002. Human Capital Inequality and Economic Growth: Some New Evidence. *The Economic Journal,* **1128**, C187–C200.

CHECCHI, DANIELE. 2000. *Does Educational Achievement Help to Explain Income Inequality?* Working Paper n.11.2000, Universitá degli Studi di Milano.

CHECCHI, DANIELE, & VAN DE WERFHORST, HERMAN G. 2017. Policies, skills and earnings: how educational inequality affects earnings inequality. *Socio-Economic Review,* **0**, 1–24.

CRESPO-CUARESMA, JESUS, SAUER, PETRA, & KC, SAMIR. 2013. *Gini Coefficients of Educationals Attainment: Age Group Specific Trends in Educational (In)equality.* WWWforEurope Working Paper no 6.

CRUCES, GUILLERMO, DOMENCH, CAROLINA GARCÍA, & GASPARINI, LEONARDO. 2011. *Inequality in Education.* UNU-WIDER Working Paper No. 2011/93.

FAN, XIBO, THOMAS, VINOD, & WANG, YAN. 2001. *Measuring Education Inequality: Gini Coefficients of Education.* World Bank Working Paper 2525.

FAN, XIBO, THOMAS, VINOD, & WANG, YAN. 2002. *A New Dataset on Inequality in Education: Gini an Theil Indices of Schooling for 140 Countries, 1960-2000.* Mimeo, The World Bank.

GOUJON, ANNE, KC, SAMIR, SPERINGER, MARKUS, BARAKAT, BILAL, POTANCOOKOVÁ, MICHAELA, EDER, JAKOB, STRIESSNIG, ERICH, BAUER, RAMON, &

LUTZ, WOLFGANG. 2016. A Harmonized Dataset on Global Educational Attainment between 1970 and 2060 - An Analytical Window into Recent Trends and Future Prospects in Human Capital Development. *Journal of Demographic Economics*, **82**, 315–363.

INTER-AMERICAN DEVELOPMENT BANK. 1999. *Facing up to Inequality in Latin America: Economic and Social Progress in Latin America, 1998-99 Report.* Johns Hopkins University Press.

LAM, DAVID, & LEVISON, DEBORAH. 1991. Declining Inequality in Schooling in Brazil and its Effects on Inequality in Earnings. *Journal of Development Economics*, **37**(1-2), 199–225.

LONDONO, JUAN LUIS. 1990. *Kuznetsian Tales with Attention to Human Capital.*

LÓPEZ, RAMÓN, THOMAS, VINOD, & WANG, YAN. 1998. *Addressing the Education Puzzle: The Distribution of Education and Economic Reform.* World Bank Working Paper 2031.

MAAS, JACOB, & CRIEL, CEERT. 1982. *Distribution of Primary School Enrollments in Eastern Africa.* Wold Bank Staff Working Paper 511.

MESCHI, ELENA, & SCERVINI, FRANCESCO. 2013. Expansion of schooling and educational inequality in Europe: the educational Kuznets curve revisited. *Oxford Economic Papers*, **66**, 660–680.

MESCHI, ELENA, & SCERVINI, FRANCESCO. 2014. A new dataset on educational inequality. *Empirical Economics*, **47**, 695–716.

MORRISSON, CHRISTIAN, & MURTIN, FABRICE. 2013. The Kuznets Curve of Human Capital Inequality: 1870-2010. *Journal of Economic Inequality*, **11**(3), 238–301.

POTANCOKOVÁ, MICHAELA, KC, SAMIR, & GOUJON, ANNE. 2014. *Global Estimates of Mean Years of Schlooling: A New Methodology.* IIASA Interim Report IR-14-005.

RAM, RATI. 1990. Educational Expansion and Schooling Inequality: International Evidence and Some Implications. *The Review of Economics and Statistics*, **72**(2), 266–274.

ROSTHAL, RICHARD A. 1978. *Measures of Disparity: A Note.*

SAUER, PETRA, & ZAGLER, MARTIN. 2012. Economic Growth and the Quantity and Distribution of Education: A Survey. *Journal of Economic Surveys*, **26(5)**, 933–951.

SAUER, PETRA, & ZAGLER, MARTIN. 2014. (In)equality in Education and Economic Development. *Review of Income and Wealth*, **60**, 353–79.

SHERET, MICHAEL. 1988. Equality Trends and Comparisons for the Education System of Papua New Guinea. *Studies in Educational Evaluation*, **14**(1), 91–112.

SILBER, JACQUES. 1989. Factor Components, Population Subgroups and the Computation of the Gini Index of Inequality. *The Review of Economics and Statistics*, **71(1)**, 107–115.

A Between-group Inequality

The contribution of inequality between any two categories or groups p, q can be written as follows (Silber, 1989),

$$
\begin{aligned}
I_{Bpq}^{E} &= -\Big(\frac{n_q}{n}\Big)n_p\bar{s}_p + \Big(\frac{n_p}{n}\Big)n_q\bar{s}_q \\
&= \frac{n_p + n_q}{n}(n_p\bar{s}_p + n_q\bar{s}_q)I_{pq}^{E}
\end{aligned}
\tag{17}
$$

where

$$
I_{pq}^{E} = \frac{n_q}{n_p + n_q}\frac{n_p\bar{s}_p}{n_p\bar{s}_p + n_q\bar{s}_q} - \frac{n_p}{n_p + n_q}\frac{n_q\bar{s}_q}{n_q\bar{s}_p + n_q\bar{s}_q}
\tag{18}
$$